RECEIVED
CENTRAL FAX CENTER
SEP 05 2007

HEWLETT-PACKARD COMPANY
Intellectual Property Administration
P.O. Box 272400
Fort Collins, Colorado 80527-2400

PATENT APPLICATION

ATTORNEY DOCKET NO. 200308713-1

IN THE
UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s): George H. Forman et al.

Confirmation No.: 8407

Application No.: 10/733,750

Examiner: Michael Le

Filing Date: December 11, 2003

Group Art Unit: 2163

Title: Iteratively Cleaning Data Records Based on Matching the Data Records to Stored Records

Mail Stop Appeal Brief-Patents
Commissioner For Patents
PO Box 1450
Alexandria, VA 22313-1450

TRANSMITTAL OF APPEAL BRIEF

Transmitted herewith is the Appeal Brief in this application with respect to the Notice of Appeal filed on July 5, 2007.

The fee for filing this Appeal Brief is (37 CFR 1.17(c)) $500.00.

(complete (a) or (b) as applicable)

The proceedings herein are for a patent application and the provisions of 37 CFR 1.136(a) apply.

☐ (a) Applicant petitions for an extension of time under 37 CFR 1.136 (fees: 37 CFR 1.17(a)-(d)) for the total number of months checked below:

☐ 1st Month $120	☐ 2nd Month $450	☐ 3rd Month $1020	☐ 4th Month $1590

☐ The extension fee has already been filed in this application.

☒ (b) Applicant believes that no extension of time is required. However, this conditional petition is being made to provide for the possibility that applicant has inadvertently overlooked the need for a petition and fee for extension of time.

Please charge to Deposit Account 08-2025 the sum of $ 500 . At any time during the pendency of this application, please charge any fees required or credit any over payment to Deposit Account 08-2025 pursuant to 37 CFR 1.25. Additionally please charge any fees to Deposit Account 08-2025 under 37 CFR 1.16 through 1.21 inclusive, and any other sections in Title 37 of the Code of Federal Regulations that may regulate fees.

☒ A duplicate copy of this transmittal letter is enclosed.

☐ I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, Alexandria, VA 22313-1450
Date of Deposit:

OR

☒ I hereby certify that this paper is being transmitted to the Patent and Trademark Office facsimile number (571)273-8300.
Date of facsimile: September 5, 2007
Typed Name: Ginger Yount
Signature: September 5, 2007

Respectfully submitted,
George H. Forman et al.
By _____
Dan C. Hu
Attorney/Agent for Applicant(s)
Reg No. : 40,025
Date : September 5, 2007
Telephone : (713) 468-8880, ext. 304

Rev 10/06a (ApiBrief)

RECEIVED
CENTRAL FAX CENTER
SEP 05 2007

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants:	George H. Forman et al.	§	Art Unit:	2163
Serial No.:	10/733,750	§	Examiner:	Michael Le
Filed:	December 11, 2003	§		
For:	Iteratively Cleaning Data Records Based on Matching the Data Records to Stored Records	§	Atty. Dkt. No.:	200308713-1 (HPC.0395US)

Mail Stop Appeal Brief-Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

APPEAL BRIEF PURSUANT TO 37 C.F.R § 41.37

Sir:

The final rejection of claims 1, 4-13, 15, 17-21, and 23 is hereby appealed.

I. REAL PARTY IN INTEREST

The real party in interest is Hewlett-Packard Development Company, L.P.

II. RELATED APPEALS AND INTERFERENCES

None.

III. STATUS OF THE CLAIMS

Claims 1, 4-13, 15, 17-21, and 23 have been finally rejected and are the subject of this appeal. Claims 2, 3, 14, 16, and 22 have been cancelled.

09/06/2007 SSITHIB1 00000037 082025 10733750
01 FC:1402 500.00 DA

Date of Deposit: September 5, 2007

I hereby certify that this correspondence is being facsimile transmitted to the U.S. Patent Office (Fax No. (571) 273-8300) on the date indicated above.

Ginger Yount

Ginger Yount

RECEIVED
CENTRAL FAX CENTER
SEP 05 2007

IV. STATUS OF AMENDMENTS

An Amendment under 37 C.F.R. § 1.116 was submitted. In the Amendment, claims 2 and 14 were cancelled to render the rejection of those claims moot, and the Abstract was amended to address a formality objection. Entry of the amendment is proper since the amendment removes issues from appeal.

V. SUMMARY OF THE CLAIMED SUBJECT MATTER

The following provides a concise explanation of the subject matter defined in each of the independent claims involved in the appeal, referring to the specification by page and line number and to the drawings by reference characters, as required by 37 C.F.R. § 41.37(c)(1)(v). Each element of the claims is identified by a corresponding reference to the specification and drawings where applicable. Note that the citation to passages in the specification and drawings for each claim element does not imply that the limitations from the specification and drawings should be read into the corresponding claim element.

Independent claim 1 recites a heuristics analysis tool embodied in a computer-readable storage medium, comprising:

> a persistent table (Fig. 1:109), having clean data records (Fig. 1:113) and key records (Fig. 1:111) wherein at least one key record is associated with each clean data record, each key record having at least one field of data from the associated clean data record (Spec., ¶ [0015]); and
>
> heuristic-based routines (Fig. 2:203) to match newly received data records (Fig. 1:103, 117) to the key records in the persistent table, the heuristic-based routines to iteratively clean the newly received data records by modifying the newly received data records in response to no match occurring between the received data records and the key records in the persistent table (Spec., ¶¶ [0014], [0016]-[0018], [0021],[0022], [0033],[0034]).

Independent claim 9 recites a data association and cleaning method comprising:

storing a plurality of clean data files (Fig. 1:113) and, associated with each of said clean data files, at least one indexing record (Fig. 1:111), each said indexing record containing at least one field related to a respective associated clean data file such that said at least one indexing record serves as a pointer to the respective associated said clean data file (Spec., ¶ [0015]);

comparing (Fig. 1:121) an input data record (Fig. 1:117) to the indexing records for obtaining a match, and if the match occurs, assigning said input data record to the respective associated said clean data file (Spec., ¶ [0016]);

if the match does not occur, iteratively cleaning (Fig. 1:125, 131) the input data record until at least a near-match between said cleaned input data record and at least one of the indexing records is obtained and assigning said cleaned input data record to the one of said clean data files associated with the near-matched indexing record (Spec., ¶¶ [0017], [0018], [0021], [0022], [0033],[0034]); and

upon a near match (Fig. 1:133), adding said cleaned input data record as a new indexing record for the associated one of said clean data files, and upon no match, adding said cleaned input data record as a new clean data file with an associated indexing record therefor (Spec., ¶ [0022]).

Independent claim 15 recites a computer memory comprising:

computer code means for receiving an input data record (Fig. 1:117; Spec., ¶ [0016]);

computer code means for comparing said input data record to a tabular format set of crude keys (Spec., ¶ [0016]);

computer code means for returning a clean key associated with one of said crude keys upon a comparing match (Spec., ¶ [0016]);

computer code means for iterative cleaning of said input data record upon a no-match return and storing the iteratively-generated respective cleaned input data record therefrom (Spec., ¶¶ [0017], [0018], [0021], [0022], [0033],[0034]);

computer code means for re-comparing said iteratively-generated respective cleaned data record to said set of crude keys (Spec., ¶ [0018]); and

computer code means for creating a new crude key from a last said iteratively-generated respective cleaned input data record such that said new crude key is added to the set of crude keys (Spec., ¶ [0018]).

Independent claim 23 recites a method of doing business comprising:

storing a database of clean data files (Fig. 1:113) for each of a plurality of entities (Spec., ¶ [0015]);

creating a tabulation (Fig. 1:109) of crude keys, each having a pointer to an associated one of said clean data files (Spec., ¶ [0015]);

receiving a dirty data record (Fig. 1:117) related to at least one entity of said plurality of entities (Spec., ¶ [0016]);

comparing (Fig. 1:121) said dirty data record to said tabulation (Spec., ¶ [0016]);

assigning said dirty data record to one of said clean data files if a match is found based on the comparing (Spec., ¶ [0016]);

cleaning (Fig. 1:125) the dirty data record by modifying the dirty data record in response to determining that no match is present based on the comparing (Spec., ¶ [0017]); and

comparing (Fig. 1:127) the cleaned dirty data record to said tabulation (Spec., ¶ [0018]).

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL[1]

A. Claims 1, 4, 5, 7, 8, 15, 17-19, 21, And 23 Rejected Under 35 U.S.C. § 103 Over U.S. Patent No. 5,819,291 (Haimowitz) In View Of U.S. Patent Application Publication No. 2002/0069195 (Commons).

B. Claim 6 Rejected Under 35 U.S.C. § 103 Over Haimowitz In View Of Commons And U.S. Patent No. 5,806,058 (Mori).

C. Claim 9-13 Rejected Under 35 U.S.C. § 103 Over Haimowitz In View Of Commons And U.S. Patent No. 5,276,616 (Kuga).

D. Claim 20 Rejected Under 35 U.S.C. § 103 Over Haimowitz In View Of Commons And U.S. Patent No. 6,070,164 (Vagnozzi).

[1] Since claims 2 and 14 have been cancelled in the Amendment under 37 C.F.R. § 1.116 that was submitted, the rejections under 35 U.S.C. § 112, ¶ 2, of those claims have been rendered moot.

RECEIVED
CENTRAL FAX CENTER
SEP 05 2007

VII. ARGUMENT

The claims do not stand or fall together. Instead, Appellant presents separate arguments for various independent and dependent claims. Each of these arguments is separately argued below and presented with separate headings and sub-headings as required by 37 C.F.R. § 41.37(c)(1)(vii).

A. Claims 1, 4, 5, 7, 8, 15, 17-19, 21, And 23 Rejected Under 35 U.S.C. § 103 Over U.S. Patent No. 5,819,291 (Haimowitz) In View Of U.S. Patent Application Publication No. 2002/0069195 (Commons).

1. Claims 1, 4, 5, 7, and 8.

Claim 1 was rejected as being obvious over Haimowitz and Commons. It is respectfully submitted that a *prima facie* case of obviousness has not been established with respect to claim 1 for at least the following reasons: (1) no reason existed that would have prompted a person of ordinary skill in the art to combine the teachings of Haimowitz and Commons to achieve the claimed subject matter (*see KSR International Co. v. Teleflex, Inc.*, 127 S. Ct. 1727, 1741, 82 U.S.P.Q.2d 1385 (2007)); and (2) the hypothetical combination of the references does not disclose or hint at all elements of claim 1.

Point (2) is addressed first. The objective teachings of the prior art references clearly indicate that the claimed subject matter is non-obvious. To make a determination under § 103, several basic factual inquiries must be performed, including (1) determining the scope and content of the prior art; and (2) ascertaining the differences between the prior art and the claims at issue. *See Graham v. John Deere Co.*, 383 U.S. 1, 17, 148 U.S.P.Q. 459 (1965).

In making the obviousness rejection, the Examiner conceded that Haimowitz fails to disclose the following subject matter of claim 1: "the heuristic-based routines to iteratively clean the newly received data records by modifying the newly received data records in response to no

match occurring between the received data records and the key records in the persistent storage." 4/4/2007 Office Action at 4. The Examiner cited Commons as disclosing certain subject matter. *Id.*

However, the obviousness rejection is clearly defective as the passages of Commons made on pages 4 and 5 of the 4/4/2007 Office Action does not disclose or hint at the claimed subject matter conceded by the Examiner to be missing from Haimowitz. On page 4 of the 4/4/2007 Office Action, the Examiner cited to paragraphs [0058] and [0060]-[0063], along with Figs. 3A and 3B, of Commons. The cited passages refer to an iterative process of finding a matching database record relating to an unidentified DVD. As explained in the cited passages of Commons, a first search key is initially used to search for a matching record in the database. Commons, ¶ [0058]. If a match is not found using the first search key, a second search key is generated, which has less specific information than the first search key. Commons, ¶ [0060]. If no match is found using the second search key, a third search key is generated from the number of chapters and frames per chapter of a DVD title. Commons, ¶ [0061]. If no match is found using the third search key, then a fourth search key is generated using a hash code that is less unique than the hash code used in the third search key. Commons, ¶ [0062]. If the fourth search key does not produce a match, then a fifth search key is generated based on the title of the unidentified DVD. Commons, ¶ [0063].

Thus, what Commons teaches is the fact that successively less unique search keys are generated to find a matching database record. However, generating successively less unique search keys, as performed by Commons, is not the same as the subject matter of claim 1, which recites heuristic-based routines to match *newly received data records* to the key records in the persistent table, the heuristic-based routines to iteratively *clean the newly received data records*

by *modifying the newly received data records* in response to no match occurring between the received data records and the key records in the persistent table.

Note that claim 1 specifically recites a persistent table have clean data records and key records—therefore, in claim 1, the clean data records and key records are not the same items. However, the Examiner appears to have confused the search keys of Commons (which may correspond to the key records of claim 1) with the clean data records of claim 1.

Moreover, successively generating new search keys as performed in Commons is clearly different from *cleaning* the newly received data records by *modifying* the newly received data records. In Commons, the first search key is based on the total number of titles, chapters per title, and number of frames per chapter. Commons, ¶ [0058]. The second search key is generated based on non-uniquely identifying information, such as be concatenating a predetermined number of characters of the volume name and hash-coded time stamp information. Commons, ¶ [0060]. The third search key is generated from the number of chapters and frames per chapter of the first title with the largest number of chapters on the unidentified DVD. Commons, ¶ [0061]. The fourth search key uses the number of chapters and frames per chapter of the first title with the largest number of chapters on the unidentified DVD, but the hash code used in the fourth search key permits the number of frames per chapter to vary by as much as 100 frames. The fifth search key is generated based on the title of the unidentified DVD. Thus, it is clear that what Commons contemplates is the *generation* of different search keys from *different combinations* of information or using different hash coding to achieve different search keys. Generating different search keys is clearly not the same as *cleaning* a newly received data record by *modifying* the newly received data record, as recited in claim 1.

In view of the foregoing, it is clear that the hypothetical combination of Haimowitz and Commons does not disclose or hint at all elements of claim 1. A *prima facie* case of obviousness has therefore not been established for at least this reason.

Moreover, no reason existed that would have prompted a person of ordinary skill in the art to combine the teachings of Haimowitz and Commons to achieve the claimed subject matter.

Haimowitz is concerned about duplication of customer records in a large business database. Haimowitz, 1:11-12. Therefore, Haimowitz proposes a matcher that receives new customer records and uses a hash key to select a set of candidates from existing records in the database. Haimowitz, 3:13-17. The matching operation performed by Haimowitz creates a list of potential matches. Haimowitz, 3:21-22. The matcher makes a decision whether to create a new customer record in the database, update an existing record in the database, or save the new data in a pending file, based on the matching. Haimowitz, 3:25-29.

However, as conceded by the Examiner, Haimowitz does not disclose or hint at the use of any heuristic-based routine to iteratively *clean* newly received data records by *modifying* the newly received data records in response to no match occurring between the received data records and the key records in the persistent table.

Moreover, as discussed above, Commons relates to generating successively less unique search keys to find information in a database for an unidentified DVD. There is absolutely no hint provided in Commons of *cleaning* new customer records such as those received in Haimowitz for matching to a database. Generating different search keys is completely unrelated to cleaning received new customer records by modifying such customer records. Therefore, in view of the foregoing, a person of ordinary skill in the art would not have found any reason to

combine the teachings of Haimowitz and Commons to achieve the claimed subject matter. The *prima facie* case of obviousness is defective for this additional reason.

In view of the foregoing, claim 1 and its dependent claims are allowable over Haimowitz and Commons. Reversal of the final rejection of the above claims is respectfully requested.

2. Claim 23.

Independent claim 23 was also rejected as being obviousness over Haimowitz and Commons. It is respectfully submitted that a *prima facie* case of obviousness has also not been established with respect to claim 23.

Claim 23 recites comparing the dirty data record to a tabulation of crude keys that each has a pointer to an associated one of clean data files in a database, assigning the dirty data record to one of the clean data files if a match is found based on the comparing, and *cleaning* the dirty data record by *modifying* the dirty data record in response to determining that no match is present based on the comparing, and *comparing the cleaned* dirty data record to the tabulation.

As conceded by the Examiner, Haimowitz does not disclose cleaning the dirty data record by modifying the dirty data record in response to determining that no match is present based on the comparing. 4/4/2007 Office Action at 7. This concession also necessarily means that Haimowitz fails to disclose comparing the *cleaned* dirty data record to the tabulation, as recited in the last clause of claim 23.

As purportedly disclosing the claimed subject matter missing from Haimowitz, the Examiner cited Commons. *Id.* Specifically, the Examiner cited the same passages of Commons as those cited against claim 1, which refer to successive generation of less unique search keys in response to no match to a database.

For similar reasons given above with respect to claim 1, Commons does not disclose cleaning the *dirty data record* by modifying the *dirty data record* in response to determining that no match is present based on the comparing. In claim 23, "clean data files," "crude keys," and "dirty data record" are expressly recited, which indicates that these items are different items. Thus, the generation of successively less unique search keys, as performed by Commons, would not constitute cleaning the *dirty data record* recited in claim 23, since the search keys may be more appropriately considered the crude keys recited in claim 23.

Even more fundamentally, successively generating less unique search keys as performed by Commons is completely different from *cleaning* the dirty data record by *modifying* the dirty data record.

Therefore, it is clear that the hypothetical combination of Haimowitz and Commons does not disclose or hint at all elements of claim 23.

Moreover, for reasons similar to those given above with respect to claim 1, no reason existed that would have prompted a person of ordinary skill in the art to combine the teachings of Haimowitz and Commons to achieve the claimed invention. Therefore, the obviousness rejection of claim 23 is defective.

Reversal of the final rejection of the above claim is respectfully requested.

3. Claims 15, 17-19, and 21.

Independent claim 15 was also rejected as being obviousness over Haimowitz and Commons. As conceded by the Examiner, Haimowitz fails to disclose computer code means for iteratively cleaning of the input data record upon a no-match return and storing the iteratively-generated respective cleaned input data record therefrom, computer code means for re-comparing the iteratively-generated respective cleaned data record to the set of crude keys, and computer

code means for creating a new crude key from a last iteratively-generated respective cleaned input data record such that the new crude key is added to the set of crude keys. 4/4/2007 Office Action at 8.

Instead, the Examiner cited to the same passages of Commons cited against claim 1 as disclosing the subject matter missing from Haimowitz. *Id.*

The obviousness rejection is clearly defective, as successively generating less unique search keys as performed by Commons is completely different from iteratively cleaning the input data record upon a no-match return, for similar reasons as discussed above.

Moreover, note that claim 15 further recites the creation of a new crude key from "a last said iteratively-generated respective cleaned input data record such that said new crude key is added to the set of crude keys." There is absolutely no indication in Haimowitz and Commons of creating a new crude key from a last iteratively-generative respective cleaned input data record, as recited in claim 15. The generation of the search keys in Commons is produced from the information associated with an unidentified DVD; Commons clearly does not disclose modifying the information associated with the DVD from which a new crude key can be created. Therefore, this is a further basis that the hypothetical combination of Haimowitz and Commons fails to disclose or hint at all elements of claim 15.

Also, as discussed above in connection with claim 1, no reason existed that would have prompted a person of ordinary skill in the art to combine the teachings of Haimowitz and Commons to achieve the subject matter of claim 15.

Therefore, the obviousness rejection of claim 15 its dependent claims is defective. Reversal of the final rejection of the above claims is respectfully requested.

B. Claim 6 Rejected Under 35 U.S.C. § 103 Over Haimowitz In View Of Commons And U.S. Patent No. 5,806,058 (Mori).

1. Claim 6.

In view of the defection obviousness of base claim 1 over Haimowitz and Commons, it is respectfully submitted that the obviousness rejection of dependent claim 6 over Haimowitz, Commons, and Mori is also defective. Therefore, reversal of the final rejection of the above claim is respectfully requested.

C. Claim 9-13 Rejected Under 35 U.S.C. § 103 Over Haimowitz In View Of Commons And U.S. Patent No. 5,276,616 (Kuga).

1. Claims 9-13.

Independent claim 9 was rejected as being obvious over Haimowitz, Commons, and Kuga. It is respectfully submitted that a *prima facie* case of obviousness has not been established with respect to claim 9 for at least the following reasons: (1) no reason existed that would have prompted a person of ordinary skill in the art to combine the references; and (2) the hypothetical combination of the references does not teach or suggest all elements of the claim.

The Examiner conceded that Haimowitz does not disclose the following task of claim 9: if the match does not occur, iteratively cleaning the input data record until at least a near-match between the cleaned input data record and at least one of the indexing records is obtained, and assigning the cleaned input data record to the one of the cleaned data files associated with the near-matched indexing record. 4/4/2007 Office Action at 12. The Examiner also conceded that Haimowitz does not disclose the following element of claim 9: upon a near match, adding the cleaned input data record as a new indexing record for the associated one of the clean data files, and upon no match, adding the cleaned input data record as a new clean data file with an associated indexing record therefor. *Id.*

The Examiner cited Commons as disclosing the first element identified above as missing from Haimowitz, and the Examiner identified Kuga as disclosing the second element identified above as missing from Haimowitz. *Id.* at 12-13.

It is respectfully submitted that the Examiner erroneously stated that Commons discloses iteratively cleaning an input data record when a match does not occur. As discussed above in connection with the other claims, Commons progressively generates less unique search keys, which cannot constitute *cleaning the input data record* recited in claim 9. Cleaning of an input data record implies that some modification of the input data record occurs to fix some aspect of the input data record, such as to remove white space, illegal characters, and so forth. *See* Specification, ¶ [0013], lines 8-11. Therefore, because of the defective application of Commons to the claim elements, it is respectfully submitted that the obviousness rejection is defective, since the hypothetical combination of Haimowitz, Commons, and Kuga does not disclose or hint at the element of iteratively cleaning the input data record if a match does not occur.

Also, contrary to the assertion by the Examiner, Kuga fails to teach or hint at adding a cleaned input data record as a new indexing record for an associated one of the clean data files upon occurrence of a near match. The Examiner cited column 13, lines 26-45, of Kuga, as disclosing this claim element. The cited passage of Kuga refers to generating an index from incoming text when a text portion from the incoming text does not match an exact entry in a dictionary, but matches a variant that is found in the dictionary. There is no hint by Kuga of adding a *cleaned* input data record as a new indexing record for the associated one of clean data files.

In the Response to Arguments section of the 4/4/2007 Office Action, the Examiner argued that Kuga does in fact disclose the cleaning of an input data record. Specifically, the

Examiner pointed to the generation of an inflection or variant of an entry to provide a match. However, providing the inflection or variant of the entry does not provide a cleaned input data record that is added as a new indexing record. As explained in column 13 of Kuga, if an incoming text portion does not match an exact entry in a dictionary, the matching module 38 checks the "inflection" or "variant" field of the dictionary 40. Kuga, 13:25-28. An inflection/variant generator 42 generates an inflection of the standard entry of the dictionary for a variant of the spelling based on the applied information and provides the inflection/variant information to the matching module 38 for matching. Kuga, 13:33-36. In one example, Kuga notes that if "On-Kun Input" is compared with "On-Kun-Input" in the dictionary, no match would occur in a first comparison. Kuga, 13:37-40. At this time, Kuga teaches that the inflection/variant generator 42 would generate the expression "On-Kun Input" from the content of the variant field in the dictionary 40. Kuga, 13:40-43. This would then cause a match to occur, which would cause the incoming text portion "On-Kun Input" to be adopted as an entry of the index and stored in the index entry storage 24. Kuga, 13:43-45. Note that the original text "On-Kun Input" has not been cleaned; rather, it is the original text portion "On-Kun Input" that is stored.

Therefore, contrary to the assertion by the Examiner, the hypothetical combination of the references clearly does not disclose or hint at the last element of claim 9.

Also, there existed no reason that would have prompted a person of ordinary skill in the art to combine the teachings of Haimowitz, Commons, and Kuga. Haimowitz relates to matching new records to existing records in a database to avoid duplication. Commons refers to iteratively searching a DVD database by starting with a unique search key to determine whether the unique search key matches information in the database. If a match does not occur, then

Commons teaches that a non-unique search key is used. This teaching in Commons regarding matching DVD information to records starting with unique search keys and proceeding to non-unique search keys is unrelated to the teachings of Haimowitz. Haimowitz is concerned with de-duplication; therefore, there would have been no motivation to incorporate a technique that uses a non-unique search key, as taught by Commons, into Haimowitz, as that would defeat the intended purpose of Haimowitz, namely identifying and avoiding duplicate records. Also, Kuga relates to matching an input text with variants that are found in a dictionary; again, a person of ordinary skill in the art would not have been motivated to apply the technique of Kuga for the purpose of matching records to existing records in a database to avoid duplication, as taught by Haimowitz.

In view of the foregoing, claim 9 and its dependent claims are non-obvious over Haimowitz, Commons, and Kuga. Reversal of the final rejection of the above claims is respectfully requested.

D. Claim 20 Rejected Under 35 U.S.C. § 103 Over Haimowitz In View Of Commons And U.S. Patent No. 6,070,164 (Vagnozzi).

1. Claim 20.

In view of the defective obviousness rejection of base claim 15 over Haimowitz and Commons, it is respectfully submitted that the obviousness rejection of dependent claim 20 over Haimowitz, Commons, and Vagnozzi is also defective.

Reversal of the final rejection of the above claim is respectfully requested.

RECEIVED
CENTRAL FAX CENTER
SEP 05 2007

CONCLUSION

In view of the foregoing, reversal of all final rejections and allowance of all pending claims is respectfully requested.

Respectfully submitted,

Date: Sept 5, 2007

Dan C. Hu
Registration No. 40,025
TROP, PRUNER & HU, P.C.
1616 South Voss Road, Suite 750
Houston, TX 77057-2631
Telephone: (713) 468-8880
Facsimile: (713) 468-8883

VIII. APPENDIX OF APPEALED CLAIMS

The claims on appeal are:

1. A heuristics analysis tool embodied in a computer-readable storage medium, comprising:

a persistent table, having clean data records and key records wherein at least one key record is associated with each clean data record, each key record having at least one field of data from the associated clean data record; and

heuristic-based routines to match newly received data records to the key records in the persistent table, the heuristic-based routines to iteratively clean the newly received data records by modifying the newly received data records in response to no match occurring between the received data records and the key records in the persistent table.

4. The tool as set forth in claim 1 wherein each said clean data record is a completely clean data file.

5. The tool as set forth in claim 1 further comprising:

at least one column recording one or more of said heuristic-based routines that were involved in generating each of said key records.

6. The tool as set forth in claim 1 further comprising:

a time-stamp associated with each said key record in the table wherein said time-stamp is indicative of most recent use.

7. The tool as set forth in claim 1 further comprising:

special flags associated with said key records, said flags associated with specific heuristic considerations.

8. The tool as set forth in claim 7 wherein one of the special flags is a quality factor assigned to each said key record.

9. A data association and cleaning method comprising:

storing a plurality of clean data files and, associated with each of said clean data files, at least one indexing record, each said indexing record containing at least one field related to a respective associated clean data file such that said at least one indexing record serves as a pointer to the respective associated said clean data file;

comparing an input data record to the indexing records for obtaining a match, and if the match occurs, assigning said input data record to the respective associated said clean data file;

if the match does not occur, iteratively cleaning the input data record until at least a near-match between said cleaned input data record and at least one of the indexing records is obtained and assigning said cleaned input data record to the one of said clean data files associated with the near-matched indexing record; and

upon a near match, adding said cleaned input data record as a new indexing record for the associated one of said clean data files, and upon no match, adding said cleaned input data record as a new clean data file with an associated indexing record therefor.

10. The method as set forth in claim 9 wherein said storing is in a displayable format.

11. The method as set forth in claim 10 further comprising:

at given intervals, performing a data clean-up on a stored table in said displayable format.

12. The method as set forth in claim 9 wherein upon said adding said cleaned input data record as a new clean data file with an associated indexing record therefor, flagging said new clean data file.

13. The method as set forth in claim 9, said iteratively cleaning further comprising:

cleaning said input data record and storing a first cleaned input data record;

comparing the first cleaned input data record to said indexing records, and

upon recognizing a match therebetween, stopping said comparing, and retrieving the associated clean data file for association with said first cleaned input data record,

upon not recognizing a match therebetween, re-cleaning said first cleaned input data record, discarding said first cleaned input data record, and storing a subsequently cleaned input data record;

re-comparing the subsequently cleaned input data set to said indexing records; and

iteratively repeating said re-cleaning and re-comparing until a predetermined phase of cleaning is reached and no said match therebetween is determined, and storing the most recent re-cleaned input data record as a new clean data file.

15. A computer memory comprising:

computer code means for receiving an input data record;

computer code means for comparing said input data record to a tabular format set of crude keys;

computer code means for returning a clean key associated with one of said crude keys upon a comparing match;

computer code means for iterative cleaning of said input data record upon a no-match return and storing the iteratively-generated respective cleaned input data record therefrom;

computer code means for re-comparing said iteratively-generated respective cleaned data record to said set of crude keys; and

computer code means for creating a new crude key from a last said iteratively-generated respective cleaned input data record such that said new crude key is added to the set of crude keys.

17. The computer memory as set forth in claim 15 wherein said computer code means for generating the new crude key has heuristic routines.

18. The computer memory as set forth in claim 17 further comprising:
computer code means for displaying in said tabular format said crude keys and heuristic routines.

19. The computer memory as set forth in claim 15 wherein each of said crude keys has an associated pointer to obtain said associated clean key.

20. The computer memory as set forth in claim 19 wherein each of said crude keys points to a cleanest one of a plurality of crude keys associated with a clean data file.

21. The computer memory as set forth in claim 15 wherein said tabular format is a displayable table, further comprising:
computer code means including heuristic routines for editing said table.

23. A method of doing business comprising:
storing a database of clean data files for each of a plurality of entities;
creating a tabulation of crude keys, each having a pointer to an associated one of said clean data files;
receiving a dirty data record related to at least one entity of said plurality of entities;
comparing said dirty data record to said tabulation;
assigning said dirty data record to one of said clean data files if a match is found based on the comparing;
cleaning the dirty data record by modifying the dirty data record in response to determining that no match is present based on the comparing; and
comparing the cleaned dirty data record to said tabulation.

IX. EVIDENCE APPENDIX

None.

X. RELATED PROCEEDINGS APPENDIX

None.